Results (by Rank) for: method for node ranking in a linked database 25 documents returned

1. 79% **A fuzzy approach to accessing accident databases**
Chung, P.W.H. ; Jefferson, M. • *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies* • 09/01/98 • 2 pages (270 words) • SUMMARY
The paper is concerned with accessing information from accident databases. It discusses the

2. 72% **Skip lists in C++. (Technology Tutorial)(Tutorial)**
Whitney, Bill • *C/C++ Users Journal* • 11/01/98 • 13 pages (3310 words) • SUMMARY
If you're like me, you're always looking for an alternative data structure that not only performs admirably, but is easy to implement and understand as well.

3. 70% **Expert Network: effective and efficient learning from human decisions in text categorization and retrieval**
Yiming Yang • *SIGIR '94. Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval* • 01/01/94 • 2 pages (270 words) • SUMMARY
Expert Network (ExpNet) is our approach to automatic categorization and retrieval of natural language texts.

4. 68% **PERCEPTIONAL LINK METHOD BASED ON DYNAMIC HYPERMEDIA SYSTEM FOR DESIGN IMAGE DATABASE SYSTEM**
FUKUDA, MANABU; KATSUMOTO, MICHIAKI; SHIBATA, YOSITAKA • *Proceedings of the 29th Hawaii International Conference on System* • 01/01/96 • 2 pages (340 words) • SUMMARY
In this paper, we introduce a Dynamic Hypermedia System (DHS) for distributed design image databases that can provide simple and flexible user access capabilities based on perceptional link, so called Kansei link method.

5. 67% **Some conditions for cost efficiency in hypermedia**
Westland, J.C. • *Information Processing & Management* • 03/01/98 • 2 pages (270 words) • SUMMARY
Recent advances in multimedia and hypertext have created new opportunities for providing information to business and consumers.

6. 66% **Top Tools To Manage Your Web Site -- Here are four tools to help you keep a Web site organized and up to date. And you don't have to buy a suite of development tools to do it.**
Rick Stout • *NetGuide* • 04/01/97 • 9 pages (2700 words) • SUMMARY
Nearly every Web site authoring tool claims site maintenance as a major feature. But most of

7. 64% **A transient hypergraph-based model for data access**
Watters, C. ; Shepherd, M.A. • *ACM Transactions on Information Systems* • 04/01/90 • 2 pages (250 words) • SUMMARY
Two major methods of accessing data in current database systems are querying and browsing. The more

8. 63% **Varghese, Turner look for ways to speed up Internet.**
St. Louis Business Journal • 05/11/98 • 6 pages (1500 words) • SUMMARY
Computer scientists at Washington University in St. Louis have patented two major inventions that

9. 62% **A parallel algorithm for optimal node ranking of a binary tree**
Sung Kwon Kim • *Journal of the Korea Information Science Society* • 07/01/92 • 2 pages (160 words) • SUMMARY
The author considers the following. Let T be a tree with n nodes. One wishes to label each node v

10. 62% **Server family delivers instant information everywhere.**
Hurd, Mark ; Pechter, Rick • *AT&T Technology* • 09/01/95 • 11 pages (2780 words) • SUMMARY
In the past few years, massively parallel processing (MPP) computers have opened previously uncharted ways for large enterprises to turn raw data into strategically important information that enables knowledge workers to make better decisions.

11. 62% **Correction of a Memory Management Method for Lock-Free Data Structures (Technical rept)**
M.M. Michael ; M.L. Scott • *NTIS* • 12/01/95 • 2 pages (210 words) • SUMMARY
Memory reuse in link-based lock-free data structures requires special care. Many lock-free

12. 62% **Identification of faulty links in dynamic-routed networks**
Wang, Clark ; Schwartz, Mischa • *IEEE J SEL AREAS COMMUN* • 01/01/93 • 2 pages (150 words) • SUMMARY
In this paper, we present a maximum a posteriori method to identify faulty links in a communication network.

13. 61% **Random sampling from B/sup +/ trees**
Olken, F. ; Rotem, D. • *Proceedings of the Fifteenth International Conference on Very Large Data Bases* • 01/01/89 • 2 pages (160 words) • SUMMARY
The authors consider the design and analysis of algorithms to retrieve simple random samples from databases.

14. 60% **A database interface integrating a query language for versions**
Andonoff, E. ; Hubert, G. ; Le Parc, A. • *Advances in Databases and Information Systems. Second East European Symposium, ADBIS'98. Proceedings* • 01/01/98 • 2 pages (140 words) • SUMMARY
This paper describes an interface for querying databases integrating versions (DBiV). This

15. 60% **The effects of a dynamic word network on information retrieval**
Iwadera, T. ; Kimoto, H. • *Proceedings of the SPIE - The International Society for Optical Engineering* • 01/01/92 • 2 pages (220 words) • SUMMARY
Describes a method of learning a user's field of interest and the effects of applying this method to information retrieval.

16. 60% **Frontal face authentication using variants of dynamic link matching based on mathematical morphology**
Kotropoulos, C. ; Tefas, A. ; Pitas, I. • *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No.98CB36269)* • 01/01/98 • 2 pages (190 words) • SUMMARY
Two variants of dynamic link matching based on mathematical morphology are developed and tested for frontal face authentication, namely, the morphological dynamic link architecture and the morphological signal decomposition-dynamic link architecture.

17. 60% **Hyperdatabase: A schema for browsing multiple databases**
M.A. Shepherd ; C.R. Watters • *NTIS* • 05/01/90 • 2 pages (270 words) • SUMMARY
In order to insure effective information retrieval, a user may need to search multiple databases on multiple systems.

18. 60% **A self-processing network model for relational databases**
De-Medonsa, E. ; Kraus, S. ; Shiftan, Y. • *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* • 04/01/99 • 2 pages (200 words) • SUMMARY
In this paper, a model which combines relational databases with self-processing networks is proposed in order to improve the performance of very large databases.

19. 60% **Perceptional link method based on dynamic hypermedia system for design image database system**
Fukuda, M. ; Katsumoto, M. ; Shibata, Y. • *Proceedings of the Twenty-Ninth Hawaii International Conference on System Sciences* • 01/01/96 • 2 pages (230 words) • SUMMARY
We introduce a dynamic hypermedia system (DHS) for distributed design image databases that can provide simple and flexible user access capabilities based on perceptional link, so called Kansei link method.

20. 60% **Bridge model: an integrated database model for office information systems**
Ozawa, H. ; Anzai, Y. ; Aiso, H. • *Transactions of the Information Processing Society of Japan* • 01/01/92 • 2 pages (150 words) • SUMMARY
Discusses dynamic and static connections within relational databases and the facilities of a link icon in the hypertext.

21. 60% **Helping the user to select a link**
Tomek, I. ; Maurer, H. • *Hypermedia* • 01/01/92 • 2 pages (170 words) • SUMMARY
Links are among the distinguishing features of hypermedia and much research resolves around them.

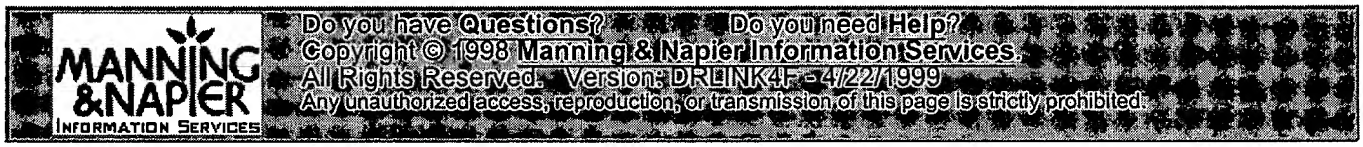
22. 59% **Rising Relevance in Search Engines.**
Notess, Greg R. • *Online* • 05/01/99 • 9 pages (2900 words) • SUMMARY
Back in the medieval days of the Internet, when a search used engine was still the software used to access bibliographic or other databases with no connection to the Internet, there was some fascinating research on statistical algorithms for sorting the output of a full-text search by projected relevance.

23. 59% **Ready for prime time?(Microsoft Windows NT operating system) (Product Development)**
Stiglich, George • *Telephony* • 07/27/98 • 7 pages (1800 words) • SUMMARY
Are Microsoft Windows NT server-based computers ready for prime-time deployment in intelligent network systems?

24. 59% **Using Informix DataBlades to facilitate E-commerce.(includes related article on executive summary) (Product Support)(Tutorial)**
Lasater, Bo • *Databased Web Advisor* • 03/01/98 • 14 pages (3600 words) • SUMMARY
Orchestrate all the capabilities an e-commerce site requires into a single, coherent, manageable system.

25. 59% **A hypermedia-based design image database system using a perceptional link method**
Shibata, Y. ; Fukuda, M. ; Katsumoto, M. • *Journal of Management Information Systems* • 12/01/96 • 2 pages (280 words) • SUMMARY
The authors introduce a hypermedia-based distributed design image database system that can provide simple and flexible user access capabilities based on the "kansei" link method.

Modify Save Alert Sort: % Rank 1/1 12/31 Newest 12/31 1/1 Oldest Source Subject Draw: Graph Tech BarChart Print... Similar Docs


MANNING
&NAPIER
INFORMATION SERVICES
Do you have Questions? Do you need Help?
Copyright © 1998 Manning & Napier Information Services.
All Rights Reserved. Version: DRLINK4F - 4/22/1999
Any unauthorized access, reproduction, or transmission of this page is strictly prohibited.

A parallel algorithm for optimal node ranking of a binary tree
Sung Kwon Kim • *Journal of the Korea Information Science Society*
Vol: 19 Issue: 4 Page: 394-9 • 07/01/92

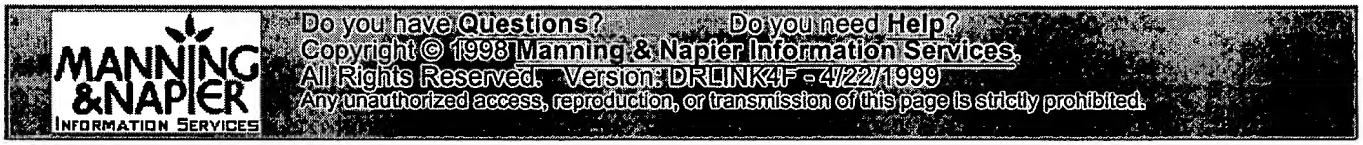
Most Relevant Section
Document Citation

The author considers the following. Let T be a tree with n **nodes**. One wishes to label each **node** v of T with a non-negative integer, **RANK**(v), so that for any two **nodes** u, v with **RANK**(u)=**RANK**(v) there must be another **node** x with **RANK**(x)<**RANK**(v) on the path between them. Such a labeling is called a **node ranking** of T. Many different **node** rankings are possible for T; among them, one which minimizes the maximum label used is called an optimal **node ranking** of T. He presents a parallel algorithm for finding an optimal **node ranking** of T when T is a binary tree. It runs in O(log n) time using n processors on the CREW PRAM.

Additional Information:

Descriptors: computational complexity ; parallel algorithms ; trees (mathematics)
Identifiers: time complexity ; parallel algorithm ; optimal **node ranking** ; binary tree ; labeling ; CREW PRAM
Document Type: Journal Paper
Language: Korean
Number of References: 6
Country of Publication: South Korea
International Standard Serial Number: 0258-9125

Document Rank: 9
Headline/Title: A parallel algorithm for optimal node ranking of a binary tree
Author(s): Sung Kwon Kim
Date: 07/01/92
Source: Journal of the Korea Information Science Society
Volume/Issue/Pg: Vol: 19 Issue: 4 Page: 394-9
Database: IEE/INSPEC
Num. Pages: 2 (160 words)
Accession Num: 4293368
MNIS Document: 137822


MANNING
&NAPIER
INFORMATION SERVICES
Do you have Questions?
Do you need Help?
Copyright © 1998 Manning & Napier Information Services.
All Rights Reserved. Version: DRLINK4F - 4/22/1999
Any unauthorized access, reproduction, or transmission of this page is strictly prohibited.

PERCEPTIONAL LINK METHOD BASED ON DYNAMIC HYPERMEDIA SYSTEM FOR DESIGN IMAGE DATABASE SYSTEM

Most Relevant Section
Document Citation

FUKUDA, MANABU; KATSUMOTO, MICHIAKI; SHIBATA, YOSITAKA • *Proceedings of the 29th Hawaii International Conference on System* • 01/01/96

In this paper, we introduce a Dynamic Hypermedia System (DHS) for distributed design image **databases** that can provide simple and flexible user access capabilities based on perceptional **link,** so called Kansei **link method**. As a proof of concept, we have developed a prototype system incorporating the DHS model. Dubbed the Textile Design Image **Database** System, this **database** aids designers using apparel CAD systems in different locations, collaborating or working separately, in the design of clothes, including kimonos. Our purpose has been to create a **database** that will allow each designer to make the best use of his or her creativity and originality- his or her (IR) style and sensitivity to beauty, (IS) J or Kansei in Japanese.

In our DHS, Metanodes are defined as abstract **nodes** and Metalinks are defined as flexible Kansei **links** respectively. Metanodes and Metalinks are combined to organize a dynamic hypermedia space from which users can easily retrieve desired design image objects by querying a knowledge agent. The knowledge agent, utilizing the knowledge-base, sets up **links** from Kansei word objects provided by the user to suitable design image objects among the multimedia **databases** distributed over the network. The knowledge agent also performs query conversion of individual users (IU) J subjective Kansei (unique, subjective use of Kansei words) into objective Kansei words using each users (IU) J individual (IR) Juser model. (IS) J These objective Kansei words are then converted to equivalent color values. Color value is the means by which all stored design images are characterized. This dynamic **linking** of Kansei word objects to equivalent design images allows individual users (IU) J Kansei to influence the retrieval process. The sophisticated and flexible CAD Systems of the future will require multimedia **database** systems with cooperative supporting capabilities similar to those of our Kansei system. (author)

Additional Information:

Keywords:
Sciences (HICSS-29)

Document Rank: 4

Headline/Title: PERCEPTIONAL LINK METHOD BASED ON DYNAMIC HYPERMEDIA SYSTEM FOR DESIGN IMAGE DATABASE SYSTEM

Author(s): FUKUDA, MANABU; KATSUMOTO, MICHIAKI; SHIBATA, YOSITAKA

Date: 01/01/96

Source: Proceedings of the 29th Hawaii International Conference on System

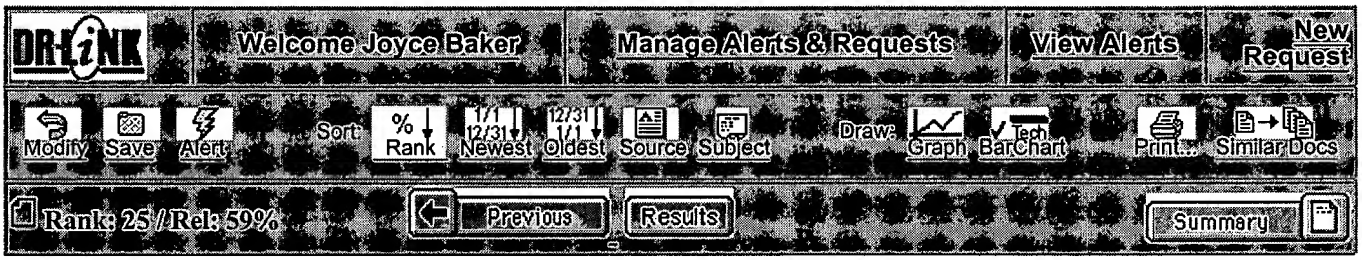
Database: Miscellaneous Software Abstracts

Num. Pages: 2 (340 words)

Accession Num: 43075

MNIS Document: 9577

Copyright Notice: © 1996 Proceedings of the 29th Hawaii International Conference on System



A hypermedia-based design image database system using a perceptional link method

Most Relevant Section
Document Citation

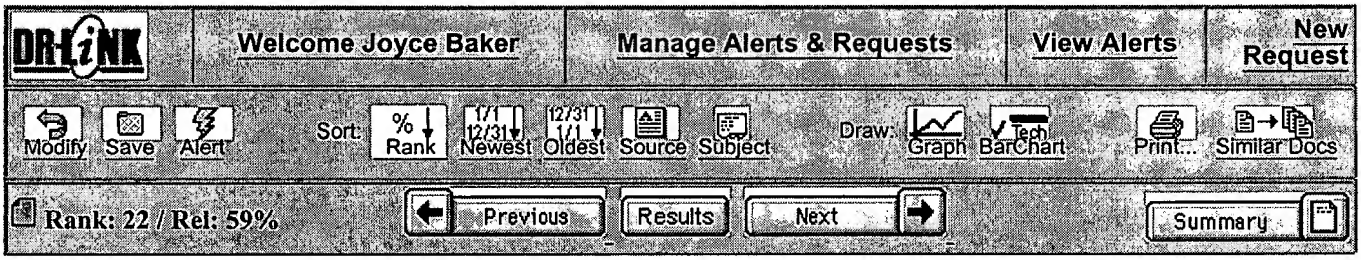
Shibata, Y. ; Fukuda, M. ; Katsumoto, M. • *Journal of Management Information Systems Vol: 13 Issue: 3 Page: 25-43* • 12/01/96

The authors introduce a hypermedia-based distributed design image **database** system that can provide simple and flexible user access capabilities based on the "kansei" **link method**. As proof of this concept, they have developed a prototype distributed multimedia information network incorporating the DHS model. Dubbed the Textile Design Image **Database** System (TDIDS), this **database** aids designers using apparel computer-aided design (CAD) systems in different locations, collaborating or working separately, in the design of clothes, including kimonos. Their purpose has been to create a **database** that will allow each designer to make the best use of his or her creativity and originality-his or her "style and sensitivity to beauty", or, in Japanese, kansei. In the hypermedia system, "metanodes" are defined as abstract **nodes** that are dynamically organized by multimedia objects, while "metalinks" are defined as flexible kansei **links**. Metanodes and metalinks are combined to organize a dynamic hypermedia space from which users can easily retrieve desired design image objects by querying a knowledge agent. The knowledge agent, utilizing the knowledge base, creates **links** from kansei word objects provided by the user to suitable design image objects among those stored on multimedia **databases** distributed across the network. The knowledge agent also performs query conversion of individual users' subjective kansei (idiosyncratic, subjective use of kansei words) into objective kansei words using each user's own user model.

Additional Information:

Descriptors: CAD ; colour ; distributed **databases** ; hypermedia ; multimedia computing ; query processing ; textile industry ; visual **databases**
Identifiers: color values ; perceptional **link method** ; hypermedia-based distributed design image **database** system ; user access ; kansei **link method** ; distributed multimedia information network ; Textile Design Image **Database** System ; apparel CAD systems ; clothes design ; creativity ; originality ; metanodes ; abstract **nodes** ; multimedia objects ; metalinks ; dynamic hypermedia space ; design image object retrieval ; knowledge agent querying ; query conversion ; subjective kansei ; objective kansei words ; user model
Document Type: Journal Paper
Number of References: 9
Author Affiliation: Toyo Univ., Saitama, Japan
Country of Publication: USA
Publisher: M.E. Sharpe
International Standard Serial Number: 0742-1222

Document Rank: 25
Headline/Title: A hypermedia-based design image database system using a perceptional link method
Author(s): Shibata, Y. ; Fukuda, M. ; Katsumoto, M.
Date: 12/01/96
Source: Journal of Management Information Systems
Volume/Issue/Pg: Vol: 13 Issue: 3 Page: 25-43
Database: IEE/INSPEC
Num. Pages: 2 (280 words)
Accession Num: 5569160
MNIS Document: 252099
Copyright Notice: © 1996 Journal of Management Information Systems



Rising Relevance in Search Engines.

Most Relevant Section
Document Citation

Notess, Greg R. • *Online Vol: 23 Issue: 3* • 05/01/99

Back in the medieval days of the Internet, when a search used engine was still the software used to access bibliographic or other **databases** with no connection to the Internet, there was some fascinating research on statistical algorithms for sorting the output of a full-text search by projected relevance. In the bibliographic realm, search output is typically sorted in reverse chronological order. Other available sort options might include an alphabetical arrangement by a specific field, such as author, title, or publication.

Efforts at sorting by relevance developed more sophistication even as the Internet moved from its computer science and defense industry roots into the popular consciousness. In some cases, output **ranked** by relevance score proved to be quite effective in research settings. Thus, in the early days of development of the Web search engines, the preferred output, if not the only seemingly sensible output, relied on relevance scores. The standard sorts used by bibliographic **databases** were not helpful with the mass of Web pages. Why sort by date when all the Web pages had been written in the past year? The HTML title element was (and often continues to be) too inconsistently used to make an alphabetical title sort very meaningful. Most Web pages had and continue to have no fielded author designation for use in sorting.

On the other hand, sorting by 1%RLs or domain name was certainly a possibility. Yet once again, in the early days of Web development, the unofficial standard of www.name.com was just beginning. Most Web pages were on sites with less meaningful names, such as xxx.lanl.gov or l2vbiol.stateu.edu or physik.technik.ch or just an IP number.

Therefore, in the early days of the Web search engines, only one option made sense, and that was relevance **ranking**. Throw a search word or two at these mammoth indexes containing words published on Web pages, and the list of hits would all be sorted by their relevance "score." The scores intended to represent how relevant the hit was to your search.

The idea was excellent. Since **databases** are so large that many searches result in thousands, if not millions of hits, just deliver the most relevant pages first--no one would be expected to manually browse millions of hits. Instead, they would only look at the first few displayed.

Unfortunately, with the disparate nature of Web pages, wide variations in file sizes, and a complete spectrum of subjects, both scholarly and mundane, determining relevance automatically is no easy task. On some searches, these early Web search engines worked successfully, providing **links** to pages that met or came close to meeting the searchers' information needs. On other searches, the relevant hits were buried deep with low relevance scores.

STANDARD RELEVANCE

The precise **methods** that each search engine uses for determining the relevance score (and thus the **ranking**) are closely guarded trade secrets. However, some general principles are discussed in their documentation or are obvious from search results.

Term frequency, positioning, weighting, and proximity are all common **ranking** criteria. The frequency of a term can be considered in several ways. Pages that have the term many

times **rank** higher, but using only this approach may artificially raise the **ranking** of very long pages that contain many words. This is sometimes evident on Web search engines when a very long page, such as a log file, is **ranked** high. A more helpful approach is where the frequency of the term is compared to the total number of words on the page.

Term positioning also certainly has a role. When a search term is found in certain sections of a Web page, it is considered more important. For example, various search engines will increase the relevance of a page if a term is found in one or more of the following areas: title, the meta keywords, meta description, first header, or first paragraph. Some search engines will ignore some of these areas while others place a larger emphasis on them. For example, Excite ignores terms in the metatags.

Term weighting refers to the practice of making some words more important than others. Infrequently used terms that do occur on certain pages would get more weight than more common terms on those pages. Stopwords are terms that receive no weight. Even on search engines that do not have stopwords, very common words will likely have a very low weight.

On searches that use more than one word, the proximity of the search terms to each other will affect the relevance scores. At a basic level, the closer the search terms are to each other, the more relevant the Web page is considered to be.

THE MISSING RELEVANCE CRITERION

I've always thought it odd that the relevance **ranking** used by the Web search engines missed some very obvious criteria to use in their relevance **ranking**. After all, when a searcher simply enters a term, such as microsoft or bowker or sprint, should not the most relevant Web pages be www.microsoft.com or www.bowker.com or www.sprint.com? Instead, many times these top-level corporate Web pages are buried deep in the results set.

Achieving this seems rather straightforward. Just add a rule that on single-word searches, a match on the term within the URL is **ranked** higher and a root URL **ranks** the highest. Just to try this, search the single term bowker on some of the main search engines and see which pages place it first in the list of relevant pages.

Lycos finds a page for Joe Bowker. Excite places an "Index-ward **database**" first, whatever that is. AltaVista tracks down Bowker '5 Books Out of Print page, but not the top-level page. Northern Light offers a page from a bulletin on the British Bowker-Saur site. HotBot takes its turn with a contact page from the U.S. Bowker site, but the searcher must choose the "See results from this site only" **link** to find the top-level Bowker page. Only Infoseek and Google! successfully find the main United States' Bowker Web page and deliver it as their number one search result.

THE SPAM DIMENSION

All the standard relevance techniques have fallen prey to an unexpected aspect of the very dynamic nature of the Web. Or perhaps more accurately, they have fallen prey to human nature. Since the Web search engines are so commonly used for finding information sites, Web builders are constantly trying to raise the profile of their site within the search engines.

Initially, the intent was to rely on author description and indexing, and the idea of metatags was born. The hope was that Web page builders would use metatags to insert keywords and descriptions that accurately represented the topic of their pages and their site. Then the search engines could give the words in author-supplied metatags a higher relevance weight.

Unfortunately, the economic underpinnings of the Web are all based on directing traffic to Web sites. Many less-than-scrupulous Web site builders quickly found that adding popular search words and phrases somewhere on their pages would attract more visitors.

Extraneous, irrelevant, and duplicative words would be added in the title, meta keywords, or the body. Adding the same word over and over again at the smallest size and in the same color as the background is a common trick.

The frequent attempts at spamming the indexes and their **ranking** cause the search engines to progressively change **ranking** algorithms and to develop sophisticated spam detection filters. In fact, any time that one of the major search engines thinks up a great new feature, they also need to consider whether or not it would be susceptible to index spamming.

The Web is very reactive. A search engine introduces an idea, e.g., AltaVista starts advertising its indexing of meta keywords. Then the spammers start to abuse the system. Another search engine, Excite, states that it will not index metatag keywords at all. Meanwhile, AltaVista and the other keyword indexers get busy trying to identify the spam techniques and create filters to get rid of those pages. Then the spammers find new ways to abuse the search engines. It becomes a never-ending cycle.

OTHER RELEVANCE FACTORS

With the huge variation in quality, document structure, information accuracy, and scope of the Web, it is a wonder that any relevance algorithm is sometimes successful. However, the new directions seen in the Clever Project at IBM's Almaden Research Center, the former Rankdex, and Google!, show an important factor that should be more widely adopted. Two factors weigh heavily in these **methods**: anchor text references and source authority.

The anchor text references use **links** from other pages. The anchor text refers to the words that have been hyper-**linked** to a new URL. In other words, a Web page that both mentions the publisher Bowker and offers a **link** to Bowker's Web site from the word "Bowker" has "Bowker" as the anchor text. When several or even many other Web sites all point to the same Web page from the same anchor text, the page to which they point is quite likely to be highly relevant to anyone searching on the term or terms within the anchor.

Unfortunately, using just the anchor **link** technique could rapidly fall prey to a new spamming technique. Web index spammers might just create loads of new pages that consist of unrelated anchors that point to their Web site. To avoid this, Google! adds a layer of weighting **links** from authoritative or well-known sites higher than anchors from unknown sites. Combining this source authority with the anchor text references can achieve highly relevant results.

PRACTICAL RELEVANCE AT WORK

While work on refining relevance algorithms for general searching is ongoing, the Internet search engines have been most successful at finding some rather simple, practical solutions to displaying highly relevant hits first. Rather than changing their relevance sorting, they have added new approaches on top of the general search results.

AltaVista's partnership with RealNames is a very basic example. On a search using AltaVista's simple search, terms that match records in the RealNames **database** are listed first-above and separate from the regular search results. Since RealNames records tie company names and trademarks to the appropriate business or organizational Web site, this practical approach achieves what most of the regular relevancy **ranking** algorithms lacked.

Another practical relevance approach is to provide both subject directory hits and results from the larger **database** of Web pages. In a sense, this is the approach Yahoo! has used so successfully. Since it is already a directory, a search on Yahoo! finds directory hits first, but then goes out for more results from the Web search engines. So the practical approach now provided by most search engines is to partner with a directory or to produce their own. Run a search on Excite, Infoseek, Snap, or Lycos and the first hits are from their directories.

Even on Alta Vista, there are **links** at the bottom of the page pointing to LookSmart categories.

Excite goes well beyond the directory addition approach. Search on Excite for microsoft and above both the Web page hits and the directory listings, Excite provides a **link** to Microsoft's Web site, their mail address, a recent stock quote, and **links** to recent news articles about the company. Then the directory **links** are offered followed by the actual Web search results (where the top hit is a Microsoft copyright statement page).

HotBot teamed up with Direct Hit to provide more practical relevance above their search results. For common searches, HotBot offers a **link** to the "Top 10 Most Visited Sites for..." These are results from Direct Hit, where the actual **links** selected by previous searchers that ran the same or a similar search are tallied, and then the most popular of these are displayed by HotBot.

Alta Vista has also been busy beyond their RealNames approach. Their partnership with Ask Jeeves provides single answer options to searches entered as questions, such as What is the best search engine? or What is the best search engine for kids? and Where can I find a basic explanation of the computer term search engine?

And then there are the ads. All the major search engines, except GoTo, clearly differentiate advertiser content from their search results. All, except GoTo, state that advertisers do not get higher relevance weighting than other non-advertiser pages. However, the advertisement placement certainly can make the ads fairly prominent and the choice of ads displayed can certainly be tied to the search terms used. In addition, sometimes the advertiser **links** actually include the search terms. Search bowker on Lycos, and one of the plain ads above the search results trumpets "Books about Bowker at barnesandnoble.com" while another invites you to "Search GTE Yellow Pages for Bowker."

Many times, especially when you have entered a complex search, these ad **links** that use the search terms make no sense. However, if the searcher is indeed actually looking for books about the topic, phone numbers, or CDs, these ads may well direct the searcher to a more appropriate information resource.

STANDARD RELEVANCE IMPROVEMENTS

While too many search engines still ignore the missing relevance criterion mentioned earlier, there have been some important improvements beyond the practical relevance approaches discussed in the previous section. Relevance **ranking** of the actual search results is still being adjusted and improved. Some of the companies are incorporating the anchor approach of Google!.

Alta Vista moved towards an automatic phrase recognition system in its simple search. Rather than processing a series of search terms as being automatically ORed together, Alta Vista looks for millions of commonly-used phrases. If such a phrase is identified, the search results are for the phrase rather than either term. For example, searching information Literacy, with no quotes, +, or other special operators, finds about 9,000 hits, as opposed to the 23 million that an OR operation would find or even the more than 100,000 that an AND operation would retrieve.

In addition, Alta Vista suggests more specific searches. That same information Literacy search run on the basic Alta Vista search finds Alta Vista suggesting other more specific phrases to search, such as Information Literacy Standards, NationaL Forum on Information Literacy, and information literacy skills. Note that the suggestions even include capital letters in some, to take advantage of Alta Vista's uppercase detection abilities. Also note that this feature, as well as the RealNames, Ask Jeeves, and automatic phrase recognition, is not available in the Alta Vista advanced search.

The success of search engines' determinations of relevance has been rising steadily, even if many would say that it still has a long way to go. Interestingly, some of the most successful changes in relevance display have come under what I have called "practical relevance approaches that deliver relevant hits separate from the regular results display. Relevance scores are no longer displayed on the results from most search engines, although newer ones like Google! still follow in the footsteps of their predecessors and give both a number and an iconic relevance score.

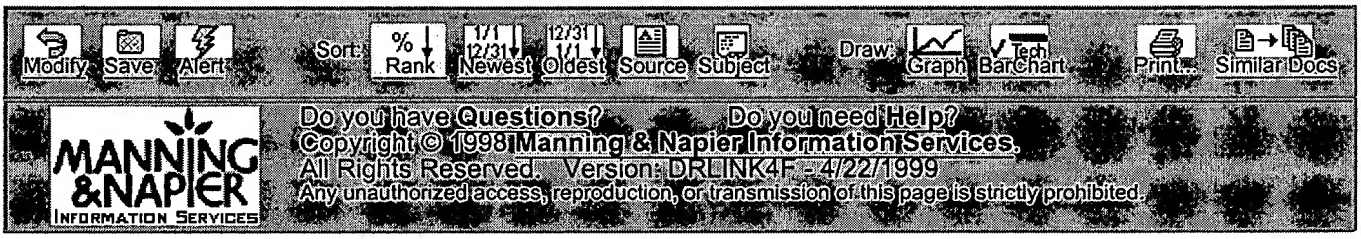
Given the very reactive and dynamic nature of the Web, and the capabilities of the Web search engines to adjust to their users needs, we can expect to see more modifications and developments. No information search and retrieval system is perfect, and the Web search engines often show some of the more obvious defects. However, even in the near term future, we can all hope to see the Web search engines delivering more relevant results more frequently and a continued rise in relevance.

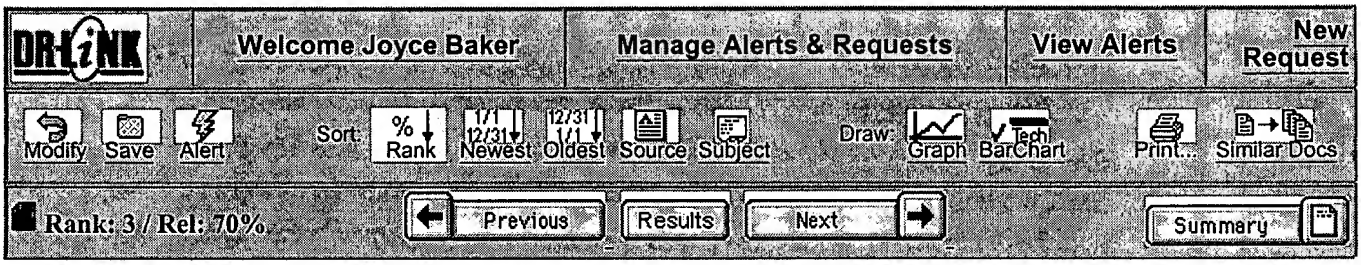
This column is also available on the ONLINE Web site at http://www.onlineinc.com / onlinemag.

Additional Information:

Audience: Trade
Publication Format: Magazine/Journal
Geographic Code/Name: 1USA United States
Industry Category: BUSN Any type of business; LIB Library and Information Science
ISSN: 0146-5422
Product Code/Name: 4811500 (Specialized Telecommunication Services)
Record Date: 19990511
Record Number: 54474837
Subject Heading: Online services--Planning
Word Count: 2606

Document Rank: 22
Headline/Title: Rising Relevance in Search Engines.
Author(s): Notess, Greg R.
Date: 05/01/99
Source: Online
Volume/Issue/Pg: Vol: 23 Issue: 3
Database: IAC Consolidated Business Collection
Num. Pages: 9 (2900 words)
Accession Num: 54474837
MNIS Document: 313038
Copyright Notice: © 1999 Online


Modify
Save
Alert
Sort:
% Rank
1/1 12/31 Newest
12/31 1/1 Oldest
Source
Subject
Draw:
Graph
Bar Chart
Print...
Similar Docs
MANNING &NAPIER
INFORMATION SERVICES
Do you have Questions?
Do you need Help?
Copyright © 1998 Manning & Napier Information Services.
All Rights Reserved. Version: DRLINK4F - 4/22/1999
Any unauthorized access, reproduction, or transmission of this page is strictly prohibited.



Expert Network: effective and efficient learning from human decisions in text categorization and retrieval

Most Relevant Section
Document Citation

Yiming Yang • *SIGIR '94. Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval Page: 13-22* • 01/01/94

Expert Network (ExpNet) is our approach to automatic categorization and retrieval of natural language texts. We use a training set of texts with expert assigned categories to construct a network which approximately reflects the conditional probabilities of categories given a text. The input **nodes** of the network are words in the training texts, the **nodes** on the intermediate level are the training texts, and the output **nodes** are categories. The **links** between **nodes** are computed based on statistics of the word distribution and the category distribution over the training set. ExpNet is used for relevance **ranking** of candidate categories of an arbitrary text in the case of text categorization, and for relevance **ranking** of documents via categories in the case of text retrieval. We have evaluated ExpNet in categorization and retrieval on a document collection of the MEDLINE **database**, and observed a performance in recall and precision comparable to the Linear Least Squares Fit (LLSF) mapping **method**, and significantly better than other **methods** tested. Computationally, ExpNet has an O(N log N) time complexity which is much more efficient than the cubic complexity of the LLSF **method**. The simplicity of the model, the high recall precision rates, and the efficient computation together make ExpNet preferable as a practical solution for real world applications.

Additional Information:

Descriptors: computational complexity ; expert systems ; information retrieval ; learning (artificial intelligence) ; natural languages ; word processing
Identifiers: Expert Network ; learning from human decisions ; text categorization ; natural language texts ; expert assigned categories ; ExpNet ; conditional probabilities ; input **nodes** ; statistics ; word distribution ; category distribution ; relevance **ranking** ; candidate categories ; text retrieval ; document collection ; MEDLINE **database** ; precision ; Linear Least Squares Fit ; LLSF mapping **method**
Document Type: Conference Paper or Conference Papers in Journal
Number of References: 15
Author Affiliation: Section of Med. Inf. Resources, Mayo Clinic, Rochester, MN, USA
Editor(s): Croft, W.B. ; van Rijsbergen, C.J.
Conference: Proceedings of 17th International Conference on Research and Development in Information Retrieval. SIGIR 94
Conference Location: Dublin, Ireland
Conference Dates: 3-6 July 1994
Place of Publication: Berlin, Germany
Country of Publication: Germany
Publisher: Springer-Verlag
Sponsoring Organization: Dublin City Univ. ; Aer Lingus ; Bord Failte ; Comm. Eur. Communities ; et al
Length of Source: 358
International Standard Book Number: 3 540 19889 X

Document Rank: 3

Headline/Title: Expert Network: effective and efficient learning from human decisions in text categorization and retrieval

Author(s): Yiming Yang

Date: 01/01/94

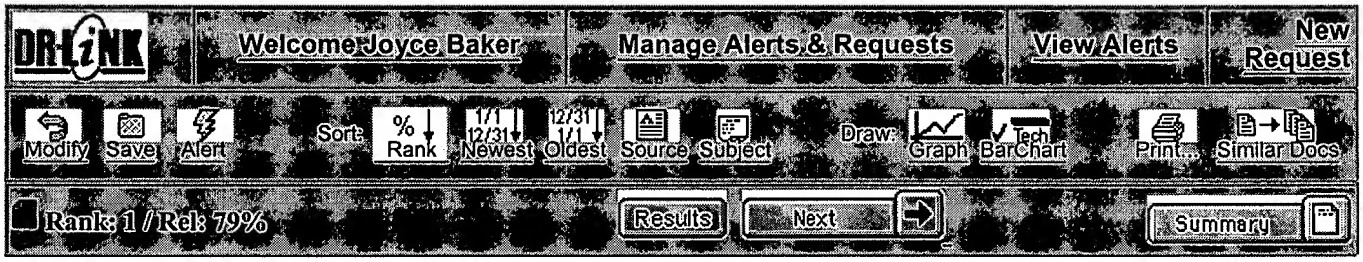
Source: SIGIR '94. Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval Page: 13-22

Database: IEE/INSPEC

Num. Pages: 2 (270 words)

Accession Num: 4910912

MNIS Document: 343813



A fuzzy approach to accessing accident databases

Most Relevant Section
Document Citation

Chung, P.W.H. ; Jefferson, M. • *Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies Vol: 9 Issue: 2 Page: 129-37* • 09/01/98

The paper is concerned with accessing information from accident **databases**. It discusses the limitation of current accident **databases** and focuses on the issue of finding and **ranking** of information that relates to a query. A user or system initiates an interaction with a **database** by specifying what is of interest in the form of a query. The query does not have to be treated as a precise description of what is of interest, but a vague or "fuzzy" one. Fuzzy **database** techniques make it possible to exploit all available information by returning not only items that match the query exactly, but also items that bear some relation to the query. A domain model for accident reports in the process industries was developed. It consists of four classification hierarchies for the attributes operation, equipment, cause and consequence. A common approach for assessing how closely two terms are related is based on the number of **links** between the two terms on a hierarchy. This approach is not appropriate for the accident **database** domain. Instead, the relationship between any two **nodes** on a hierarchy is classified into four different types. **Methods** for determining similarities for the different types of relationships are discussed and have been implemented in an accident **database**. The **ranking** of the retrieved information is much more satisfactory then the "distance" based approach.

Additional Information:

Descriptors: accidents ; **database** management systems ; fuzzy logic ; query processing
Identifiers: accident **databases** ; information **ranking** ; fuzzy **database** techniques ; attributes operation ; cause ; consequence
Document Type: Journal Paper
Number of References: 27
Author Affiliation: Dept. of Chem. Eng., Loughborough Univ. of Technol., UK
Country of Publication: Netherlands
Publisher: Kluwer Academic Publishers
International Standard Serial Number: 0924-669X

Document Rank: 1
Headline/Title: A fuzzy approach to accessing accident databases
Author(s): Chung, P.W.H. ; Jefferson, M.
Date: 09/01/98
Source: Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies
Volume/Issue/Pg: Vol: 9 Issue: 2 Page: 129-37
Database: IEE/INSPEC
Num. Pages: 2 (270 words)
Accession Num: 6066647
MNIS Document: 212868
Copyright Notice: © 1998 Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies